
Resilience indicators for rural New Zealand towns Statistical analysis

AgResearch Limited

June 2019



Private & Confidential

AgResearch Limited
Grasslands Research Centre
Tennent Drive
Palmerston North 4442

Attention: Margaret Brown

10 June 2019

Resilient Rural Communities – community data analysis

Dear Margaret

I have the pleasure of providing to you this report that extends our work on community resilience. As you know, we have been conducting case study research with a combination of qualitative and quantitative data. This report is a fully quantitative assessment of New Zealand rural areas, providing another layer of understanding regarding rural communities.

This report is provided in accordance with our contract dated 26 October 2018. Please note that it should be read in conjunction with the contract and with the restrictions in the appendix of the report.

Yours sincerely

A handwritten signature in black ink, appearing to read 'W- VJB', with a long horizontal flourish extending to the right.

Bill Kaye-Blake
Chief Economist, Wellington
Consulting
E: bill.h.kaye-blake@pwc.com
M: 021 823 000

Table of contents

Executive summary	1
The analysis extended earlier case study work	1
We analysed Census data to test our ideas of resilience	1
Statistics tell us something but not everything	1
Introduction	2
The purpose was quantitative analysis to complement case study research	2
The report builds on three years of case study research	2
This report focuses on the statistical analysis	2
Data	4
Using the Resilience Framework	4
Identifying rural towns	4
Constructing the indicators	5
Method	7
Introduction	7
Descriptive statistics	7
Correlation analysis	7
Principal component analysis	7
Results	9
Descriptive statistics	9
Correlations	9
Principal component analysis	12
Discussion	16
The PCA summarises something meaningful in the indicators	16
The analysis can be a useful initial tool for researchers	16
References	18
Appendix A Results for all CAUs	19
Appendix B Restrictions	25

Executive summary

Analysis and implications

The analysis extended earlier case study work

The AgResearch Resilient Rural Communities (RRC) programme engaged PwC to conduct analysis of official statistics for rural areas of New Zealand. The analysis extended earlier work that was based on holding workshops in rural communities. The aim was to have a larger data sample than the four original communities in the research so as to produce more statistically robust results.

The prior work found that official statistics such as Census data were somewhat reflective of the experiences of rural communities, but did not tell the whole story. Residents' perceptions of their own resilience were somewhat different from the official statistics and from expert judgments about communities' resilience.

We analysed Census data to test our ideas of resilience

We obtained data for 308 Area Units from the 2013 New Zealand Census to develop indicators for the social, economic and cultural dimensions of resilience. We analysed the indicators for patterns that would suggest whether community resilience can be identified or determined based solely on official statistics. We found the following:

- There were high levels of correlation between pairs of indicators, and there were some pairs with little correlation. The correlations suggested that there were consistent patterns that held across the resilience dimensions.
- A principal component analysis produced a single component or trend that explained about one-half of the variation in the dataset. This finding suggested that Area Units could be ordered from less resilient to more resilient. It also suggested that the ranking would miss some variation in the experience of people living in rural communities.

Statistics tell us something but not everything

The results suggest that official statistics can provide a simple indication of resilience for rural communities in New Zealand, but that this indication will not capture the full experience of these communities.

This finding could be helpful for AgResearch and the agricultural sector in a few ways:

- Researchers can quickly assess an area, such as a town or Census Area Unit, using easily available data, and have some confidence in their assessment.
- However, the finding also validates the importance of engaging with stakeholders and communities, because the view based on statistics could be incomplete.
- The statistical assessment could be useful for prioritising areas according to resilience, or for linking particular research topics to relevant locations. It could also be useful for locating anomalies – communities that seem to be performing significantly better or worse than their statistical profile. Finally, it could suggest areas of emphasis, such as resilience dimensions that could be the focus of future research.

Introduction

The purpose was quantitative analysis to complement case study research

AgResearch Limited (AgResearch) engaged PwC to conduct this research as part of the Resilient Rural Communities (RRC) programme. The purpose of the research discussed in this report was to conduct quantitative analysis that would complement case studies conducted in 2017, 2018 and 2019. The aim was to collect data on rural areas, based on the existing resilience research, and analyse it for useful patterns.

The report builds on three years of case study research

In previous years, AgResearch and PwC developed a method for measuring the resilience of rural communities. The work combined qualitative research and quantitative data analysis to test the idea of a resilience threshold for rural towns in New Zealand. The elements of the research were:

- a resilience framework that identified dimensions of resilience, with applications for researching, planning and communities (Fielke, Kaye-Blake, & Vibart, 2017; Fielke, et al., 2018)
- a workshop method that encouraged engagement and learning and could be deployed in different locations (Payne, et al., 2018)
- an analytical framework that compared three sources of information: community perceptions of resilience, official statistics and expert judgments (Kaye-Blake, Stirrat, Smith, & Fielke, 2017).

In the 2016-2017 year, RRC conducted a pilot research project to test the method. The test was successful, and demonstrated that it was possible to collect the three types of data and conduct an integrated analysis (Kaye-Blake, Stirrat, Smith, & Fielke, 2017). In 2017, researchers held a workshop with experts in the field of rural community research to decide on next steps for the research. The experts suggested that the appropriate next step was to return to the four communities in the pilot study and conduct new workshops, with the aim of increasing the number of participants and collecting additional data. In the 2017-2018 year, RRC researchers followed this advice, conducting a second round of workshops in the same four communities (Payne, et al., 2018). In 2018-2019, the focus was on responding to two ideas from the 2017 expert workshop:

- What would international comparative research demonstrate about resilience thresholds and the applicability of New Zealand findings?
- Can the research achieve a large sample size and estimate more statistically robust results?

This report addresses the second idea. It uses the concept of Macro-Social Accounting (Young & MacCannell, 1979; MacCannell, 1988) to develop a census of all the rural areas of New Zealand to investigate social, economic and cultural patterns related to community resilience. The research comprises the largest possible coverage of rural areas in New Zealand, because it includes all of them (subject to definitions as described below). It is therefore a census, not sample, and the estimates reported are for the population rather than a sample. However, it is a snapshot in time rather than longitudinal research. The analysis is therefore more about community wellbeing or its current state; the links with resilience must be inferred.

This report focuses on the statistical analysis

This report describes the statistical work undertaken. The next section describes the data that were assembled from the New Zealand Census, including definitions for the rural areas analysed. The following section describes the analytical methods used: descriptive statistics, correlation and principal component

analysis. The report then presents the results of the analysis, which led to a rank-ordered list of the Census Area Units in the data set. The final section discusses the implications of the findings.

The report focuses on the statistical analysis. The wider context of the Resilient Rural Communities programme is described elsewhere (Brown, Kaye-Blake, & Payne, 2019), and the case study research with rural communities has also been reported (Kaye-Blake, Stirrat, Smith, & Fielke, 2017; Payne, et al., 2018). This report focuses on the new contribution to the research that the statistical analysis makes.

Data

Using the Resilience Framework

The present analysis is a continuation of work on resilience in the Resilient Rural Communities programme, and specifically the community-based workshops that have been running for three years. The workshops used the Resilience Framework to organise the workshop facilitation and discussion, in order to ensure that they covered all the dimensions of resilience. In developing the framework and workshop, the resilience literature was reviewed in order to identify variables, statistics or metrics that have been used to measure resilience. The review produced a list of indicators identified in the literature and available from Statistics NZ (Fielke, Kaye-Blake, & Vibart, 2017; Payne, et al., 2018).

Table 1 Data collected for the analysis

Resilience dimension	Relevant data
Social	Population change, dependency ratio, secondary education, tertiary education, access to telephone, access to Internet, volunteering rate, vehicle ownership, smoking
Economic	Unemployment
Cultural	Religious affiliation, speaking te reo Māori, being born overseas, identifying as Māori
Institutional	Not included (state-owned housing and voter turnout were previously identified but not used here)
Environmental	Not available at CAU level

Identifying rural towns

Prior work on this topic in Resilient Rural Communities has used a case study approach. The work focused on four communities in the North Island, where two sets of workshops were held (Payne, et al., 2018). In 2019, workshops were also held in Vermont in the United States, in order to make comparisons between the two sets of communities. The results of those workshops are still being analysed. The prior research therefore had the limitation of a small sample size; only four New Zealand communities have been studied in that way.

A different approach is to collect the relevant data for all towns or communities that meet some definition and look for statistical trends in the data (MacCannell, 1988; Young & MacCannell, 1979). This approach was used for the present research. The data were taken from the 2013 New Zealand Census of Population and Dwellings. Data for the statistics in Table 1 were obtained for all Census Area Units (CAUs) in New Zealand, and then data were extracted for all small, rural CAUs. Because of the definitions of CAUs, they do not represent towns. For example, Kaitaia and Morrinsville are both divided into 'East' and 'West' CAUs. While this statistical approach makes the quantitative analysis inconsistent with the community workshops, it does reduce the amount of researcher judgment applied in developing the dataset.

For the 2013 Census, Statistics NZ used an area unit classification that divided New Zealand into 2,020 Census Area Units. CAUs are defined as 'aggregations of meshblocks. They are non-administrative areas that are in between meshblocks and territorial authorities in size' (Statistics NZ, nd). The CAUs include cities, towns, suburbs, rural areas and remote places, including the Ross Dependency and various smaller islands. The CAUs, as the first aggregation of Census meshblocks, were selected as the unit of analysis.

Two definitions were required for the analysis: ‘rural’ and ‘small’. Statistics NZ has explored the issue of classifying ‘rural’ and ‘urban’ areas in some detail (Statistics New Zealand, 2004). As a result, all CAUs are categorised according to the extent of urban influence or rural remoteness. The present analysis, however, started at the meshblock level, in order to maintain some flexibility about defining communities as rural. A meshblock ‘is the smallest geographic unit for which statistical data is reported by Stats NZ. A meshblock is a defined geographic area, varying in size from part of a city block to large areas of rural land’ (Statistics NZ, nd). All meshblocks in New Zealand are categorised into one of several categories, as shown in Table 2.

Table 2 Meshblocks by urban/rural classification

Urban/rural designation	Number of meshblocks
Area outside urban/rural profile	683
Highly rural/remote area	2,606
Independent Urban Area	5,230
Main urban area	24,618
Rural area with high urban influence	1,142
Rural area with low urban influence	3,777
Rural area with moderate urban influence	2,156
Satellite Urban Area	1,180
Total	41,392

Source for data: Statistics NZ Urban/Rural Profile classification.

For this analysis, meshblocks were considered ‘rural’ if they were Highly rural/remote areas, Independent Urban Areas, Rural areas with low urban influence or Rural areas with moderate urban influence. CAUs were then considered rural if at least 80 per cent of the constituent meshblocks fell into one of those four categories. Of the 2,020 CAUs, 613 met this criterion. For comparison, Statistics NZ listed 607 CAUs as belonging to one of those four categories in 2001.

The CAUs were also analysed to determine the ones that could be considered ‘small’. For RRC, the focus has been on rural communities with populations of between 4,500 and 10,000 people. For this analysis, the lower end was reduced to 1,000 people, particularly because CAUs do not necessarily represent whole towns. For example, the earlier RRC workshops treated Huntly as a single community while the CAU approach treats Huntly West and Huntly East as separate. All CAUs between 1,000 and 10,000 in population were therefore included in the analysis.

The filtering approach above produced a list of CAUs considered ‘rural’ and ‘small’. One CAU – Burnham Military Camp – was removed during the initial analysis because it appeared to be an outlier for the indicators. All other CAUs were retained for the final list of 308 CAUs.

Constructing the indicators

The variables that were selected for analysis required some preparation. The variables tended to be raw counts of people or households. For comparison across the CAUs, the raw counts were converted into proportions.

One issue that arose repeatedly was the correct denominator for calculating proportions. For several Census questions, the reported ‘Total’ field – the total number of people or households reported for a question – was different from the sum of the counts of the individual categories. The difference appeared to have two main sources. One source was likely the effect of confidentialising data by rounding to base three. This is a standard practice for Statistics NZ, and it has more impact on data from smaller CAUs than larger CAUs and more impact when there are many response categories (so more categories with low numbers of

responses). The other source of the difference was incomplete responses: some Census participants did not fully complete the questionnaires. The final indicators are described in Table 3 and represent our best effort to calculate them consistently across the analysis.

Table 3 Construction of indicators from Statistic NZ data

Indicator	Construction
Population change	$(\text{Census 2006 population} - \text{Census 2013 population}) \div (\text{Census 2006 population})$
Dependency ratio	$(\text{Count of people 0 to 100 years plus} - \text{Count of people 15 to 64 years of age}) \div (\text{Count of people 15 to 64 years of age})$
Secondary education	$(\text{Count of people with educational achievement from Level 2 through Doctorate, including Overseas secondary school qualification}) \div (\text{Count of people who stated an educational qualification})$
Tertiary education	$(\text{Count of people with educational achievement from Level 4 through Doctorate qualification}) \div (\text{Count of people who stated an educational qualification})$
Bachelor's degree	$(\text{Count of people with Bachelor's, Post-graduate, Masterate, or Doctorate qualification}) \div (\text{Count of people who stated an educational qualification})$
No access to telecommunications	$(\text{Count of households with No access to telecommunication systems}) \div (\text{Total households stated})$
Access to Internet	$(\text{Count of households with Access to the Internet}) \div (\text{Total households stated})$
Unemployment	$(\text{Unemployed}) \div (\text{Total people, in labour force})$
Te reo percentage	From the Languages spoken data: $(\text{Count of Maori}) \div (\text{Count of Total people stated})$ No allowance was made for 'None (eg too young to talk)'
No vehicle percentage	From the Number of motor vehicles by tenure of household: $(\text{Count of No motor vehicle}) \div (\text{Count of Total households stated})$
Maori percentage	$(\text{Count of Maori descent}) \div (\text{Count of Total people stated})$
Percent smokers	$(\text{Count of Regular smoker}) \div (\text{Count of Total people stated})$
Volunteering percentage	From the Unpaid activities (total responses) by age group and sex: $(\text{Count of Other helping or voluntary work for or through any organisation, group or marae}) \div (\text{Count of Total people stated})$
Religious affiliation	$(\text{Count of Total people, with at least one religious affiliation}) \div (\text{Count of Total people stated})$
Home ownership percentage	From the Tenure of household by number of bedrooms, for households in occupied private dwellings: $(\text{Count of Dwelling owned or partly owned plus Count of Dwelling held in a family trust}) \div (\text{Count of Total households stated})$

Method

Introduction

The data set contained 15 indicators for 308 rural CAUs in New Zealand. It was subjected to analysis using a few statistical techniques, which are described below. The analysis was conducted using the software R.¹

Descriptive statistics

The first analysis of the data was descriptive statistics. The minimum and maximum values for the indicators across the CAUs were calculated, as were the mean value and the standard deviation. These figures provided summary information about the range of values that each of the variables takes, as well as an indication of the variability across the data set.

Correlation analysis

Correlation is a measure of the association between two continuous variables and is used to determine how strongly one variable is related to or predicts a second variable (see Rodgers & Nicewander (1988) for a discussion). A common correlation statistic is Pearson's r , which generates a dimensionless index that ranges from -1 to 1. There is no probability associated with the correlation statistic, so interpretation of the statistic is by convention. Correlations greater than 0.3 in absolute value (that is, ± 0.3) are generally considered somewhat important, and those greater than 0.5 are important. The statistic provides no indication of causality. Two variables may be correlated because one causes the other, or because both are caused by something else, or merely by chance.

Correlation statistics were calculated for all pairs of the 15 variables. They can be presented as a lower triangular matrix that includes 15 diagonal elements (equal to 1.0) and 105 off-diagonal correlation coefficients.

Principal component analysis

Principal component analysis (PCA) is a statistical technique for reducing the dimensionality of a data set. There are many explanations of the technique; this discussion relies on Vyas & Kumaranayake (2006) and NIST/SEMATECH (2012) as well as linear algebra from Kolman (1986). Conceptually, the technique finds underlying or latent variables – the components – that explain or represent the variation in the data and are uncorrelated with each other. If enough principal components are estimated, then they fully represent all the variation or information in the original data set. However, the principal components will be orthogonal to each other, while variables in the original data could be correlated.

In PCA, the first component is calculated to represent ('explain') more variance than subsequent components. That is, the first component contains the most information about the data set, and each subsequent component adds successively less and less information. When using PCA to reduce the dimensionality of a data set, therefore, the first component is the most important because it represents the least loss of information. The second component is the second most important, etc.

Mathematically, PCA is a matrix transformation of a data set of correlated variables into a matrix of uncorrelated vectors. It uses the eigenvectors of the correlation matrix or co-variance matrix, depending on the technique used, and the eigenvalues of the eigenvectors. As a result, the full PCA matrix contains all the variation of the original data set, merely recalculated into the smallest possible number of orthogonal vectors (which could be less than or equal to the number of variables).

¹ R version 3.3.3 (2017-03-06) -- "Another Canoe". Copyright (C) 2017 The R Foundation for Statistical Computing Platform.

The analysis was conducted using the 'princomp' function in R.

Vyas & Kumaranayake (2006) described the use of PCA to reduce multi-dimensional socio-economic data into a linear scale as a four-part process. First, data on socio-economic information are collected. Second, the PCA is run on the data. Vyas & Kumaranayake (2006) focused on the first principal component, noting that the first principal component from a data set of socio-economic variables is assumed to represent economic status, and that prior work found the first component to explain 12 per cent to 27 per cent of total variation. The third step is interpretation. The statistical analysis produces a set of weights that indicate the contribution of each data variable to the principal component. These weights are then interpreted by researchers, using the sign (direction) and size of the weights to describe the link between the data and the factor. Fourth, the first principal component can be used to create a linear scale, which can then be used to create groups, quintiles or other classifications.

In the present research, the first step was assembling a set of indicators for the rural CAUs and their performance across the resilience dimensions. The second step was to run a PCA. The third step was to review the first and second principal components for the calculated weight of the variables. The fourth step was the calculation of the weighted score for each CAU and arrangement into a single list. The results are presented in the next section.

Results

Descriptive statistics

The descriptive statistics are reported in Table 4. The minimum, mean, maximum and standard deviation for each indicator are reported.

The summary statistics highlighted the variability of rural CAUs in New Zealand. Across all the indicators, there was a considerable gap between the minimum and maximum values. For example, population change ranged from -22 per cent (a loss of over one-fifth of the population, in Waiouru) to +167 per cent (a more-than-doubling of the population, in Lake Hayes South). Similarly, educational attainment showed large ranges. For example, the proportion with a Bachelor's degree ranged from 3 per cent to 32 per cent, a full order-of-magnitude difference. Cultural indicators displayed the same wide range. The percentage speaking te reo Māori had a minimum of nil and a maximum of nearly one-half the CAU's population. The percentage of residents of a CAU identifying as Māori ranged from four per cent to 92 per cent.

The means for the indicators provided some information on rural New Zealand as a whole. Importantly, these figures are the means for the values calculated at the CAU level, so they are not the mean of the rural population. That is, a CAU with a small population has the same weighting in these means as a CAU with a large population. Nevertheless, they provided useful information. For example, the net population change was positive over 2006 to 2013, and at four per cent was only slightly less than that growth of five per cent for the whole country. Access to telecommunications and the Internet was good on average: the mean figure for no access to telecommunication was just two per cent, and there was 70 per cent penetration of the Internet in 2013. Average unemployment was just six per cent, not far from the target unemployment for the whole economy, and home ownership was similar to national averages. The mean proportion of residents identifying as Māori was higher than the national average, which suggested that Māori were more likely to live in rural areas.

Correlations

The next analysis calculated the correlation coefficients. The results are presented in Table 5. In the table, all the correlations above $|0.3|$ (absolute value of 0.3) were shaded. The shading clearly revealed the high degree of correlation among the indicators: 78 of the 105 off-diagonal correlations were shaded. Even at a higher cut-off, there was considerable correlation: 34 correlations had an absolute value over 0.5. Some of the correlation was unsurprising: the three different education indicators (secondary education, tertiary education and Bachelor's degrees) were over 90 per cent correlated. Notably, the proportion identifying as Māori and proportion speaking te reo Māori were also over 90 per cent correlated. Other correlations were less direct. For example, the correlation between home ownership and smokers was -0.520; lower home ownership correlated with higher rates of smoking, and the correlation at the CAU level was over 50 per cent. Some correlations also suggested a level of structural difficulty or lock-in across the indicators. For example, the education indicators were correlated with access to telecommunications and the Internet, with education levels correlating with access to the Internet by over 70 per cent. This result suggests that it could be difficult to access information and educational resources without an existing base of educational attainment.

Also interesting were the indicators that did not correlate as highly across the data set. The rate of volunteering correlated well with only four of the possible 14 pairs of indicators. Home ownership correlated with only six of the other indicators and the dependency ratio correlated with only seven (half) of them. These indicators are therefore reflecting or capturing something different about these rural CAUs than the other indicators, such as education or unemployment.

Table 4 Descriptive statistics for the CAU variables

	Population change	Dependency ratio	Secondary	Tertiary	Bachelor	No access to telecommunications	Access to Internet	Unemployment	Te reo percentage	No vehicle percentage	Maori percentage	Percent smokers	Volunteering percentage	Religious affiliation	Home ownership percentage
Minimum	-0.22	0.17	0.33	0.15	0.03	0.00	0.38	0.00	0.00	0.00	0.04	0.05	0.09	0.36	0.14
Mean	0.04	0.61	0.55	0.32	0.12	0.02	0.70	0.06	0.05	0.06	0.23	0.19	0.19	0.53	0.66
Maximum	1.67	1.18	0.86	0.55	0.32	0.13	0.96	0.27	0.48	0.27	0.92	0.42	0.34	0.74	0.88
Standard deviation	0.14	0.14	0.09	0.07	0.05	0.02	0.10	0.04	0.07	0.05	0.17	0.06	0.04	0.06	0.10

Table 5 Correlation coefficients for CAU variables

	Population change	Dependency ratio	Secondary	Tertiary	Bachelor	No access to telecommunications	Access to Internet	Unemployment	Te reo percentage	No vehicle percentage	Maori percentage	Percent smokers	Volunteering percentage	Religious affiliation	Home ownership percentage
Population change	1.000														
Dependency ratio	-0.124	1.000													
Secondary	0.394	-0.448	1.000												
Tertiary	0.390	-0.318	0.953	1.000											
Bachelor	0.363	-0.384	0.931	0.937	1.000										
No access to telecommunications	-0.336	0.039	-0.364	-0.393	-0.279	1.000									
Access to Internet	0.481	-0.455	0.802	0.786	0.708	-0.677	1.000								
Unemployment	-0.330	0.410	-0.578	-0.557	-0.482	0.651	-0.780	1.000							
Te reo percentage	-0.327	0.127	-0.400	-0.403	-0.312	0.781	-0.678	0.764	1.000						
No vehicle percentage	-0.357	0.455	-0.534	-0.539	-0.458	0.521	-0.769	0.682	0.419	1.000					
Maori percentage	-0.387	0.183	-0.498	-0.498	-0.412	0.740	-0.744	0.838	0.933	0.486	1.000				
Percent smokers	-0.433	0.126	-0.726	-0.770	-0.671	0.687	-0.839	0.769	0.707	0.662	0.788	1.000			
Volunteering percentage	-0.202	0.167	-0.106	-0.048	-0.011	0.423	-0.284	0.283	0.535	0.052	0.453	0.176	1.000		
Religious affiliation	-0.273	0.397	-0.364	-0.348	-0.284	0.236	-0.424	0.334	0.410	0.371	0.357	0.251	0.399	1.000	
Home ownership percentage	0.272	0.215	0.130	0.300	0.165	-0.442	0.381	-0.287	-0.327	-0.434	-0.393	-0.520	0.016	-0.078	1.000

Note: Shaded cells are greater than 0.3 in absolute value.

Principal component analysis

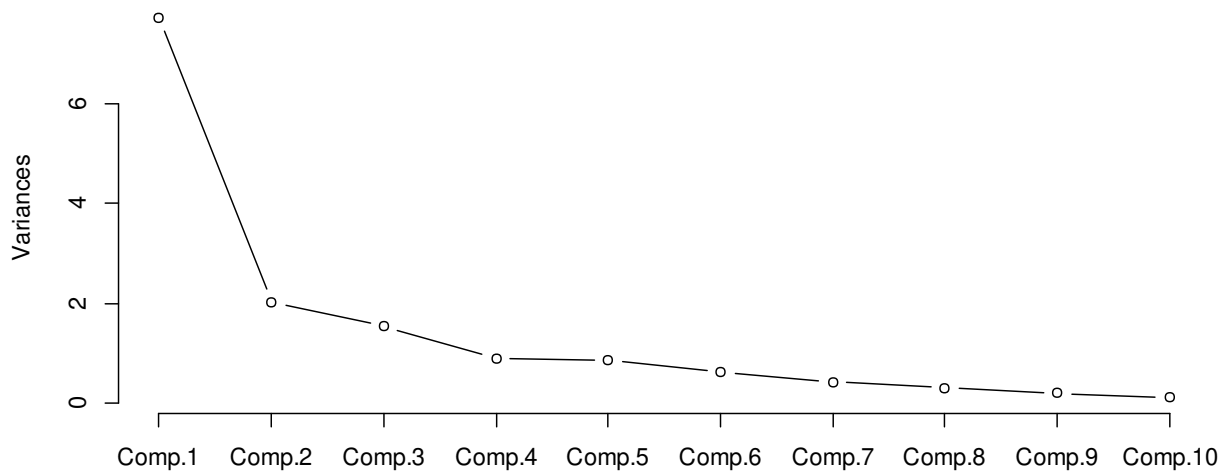
The descriptive statistics analysis found that there was considerable variation in the indicators across the CAUs. This variation is potential information for the statistical analysis to explain. The correlation analysis found that many of the indicators were highly correlated: that is, the variation to be explained was quite similar for many of the indicators. However, some of the indicators correlated poorly with the others, which means that they provided more variance and information to be explained. Principal component analysis (PCA) was a tool to summarise the variation in the data set.

The PCA estimated 15 components for the data set of 15 indicators. The proportion of variance explained by the components is presented numerically in Table 6, including the additional proportion explained by the component and the total (cumulative) proportion explained by the component and all previous components. The first component captured over half – 51.3 per cent – of the variance. The next two components each accounted for more than 10 per cent of the variation (13.5 per cent and 10.3 per cent, respectively). The final six components each captured less than one per cent of the variation, with the final component accounting for only 0.15 per cent of the total variance in the data set. The results are presented graphically in Figure 1, a scree plot for the PCA, which indicates the large difference between the first component and the others.

Table 6 Variance captured by PCA component

Component	Standard deviation	Proportion of variance	Cumulative proportion
Component 1	2.77	0.513	0.513
Component 2	1.42	0.135	0.647
Component 3	1.24	0.103	0.750
Component 4	0.95	0.060	0.810
Component 5	0.93	0.058	0.868
Component 6	0.79	0.042	0.910
Component 7	0.65	0.028	0.938
Component 8	0.57	0.021	0.960
Component 9	0.46	0.014	0.974
Component 10	0.35	0.008	0.982
Component 11	0.28	0.005	0.987
Component 12	0.28	0.005	0.993
Component 13	0.22	0.003	0.996
Component 14	0.21	0.003	0.998
Component 15	0.15	0.002	1.000

Figure 1 PCA result: variance explained by components



The loadings for all 15 principal components are given in Table 7. These are the weights that each indicator contributes to each of the principal components – how much each indicator contributes the component. Following Vyas & Kumaranayake (2006), the analysis focused on the first principal component to construct a socio-economic status scale. The first principal component has loadings for all indicators. Thus, the first principal component includes an influence from all the indicators included in the data set. Population change, educational attainment, access to telecommunications (a negatively coded indicator) and the Internet, and housing ownership all contribute positively to the component. The dependency ratio, unemployment, lack of a vehicle, proportion of smokers and rates of volunteering and religious affiliation are all negatively loaded on this component. Proportions of te reo speakers and Māori identification are also negatively loaded on the component.

These loadings were used to create a score for each of the 308 CAUs in the data set. The value of each indicator was multiplied by the corresponding loading and the result summed for each CAU. This calculation yielded a score based on the first principal component. Finally, the CAUs were reordered according to their scores. The full list is provided in the appendix.

A selection of scores is provided in Table 8. The CAUs with the 10 lowest and highest scores are listed, along with 10 CAUs from the middle of the list. Inspection of the list suggests that the scoring based on the first principal component reflected something with regard to resilience or wellbeing or socio-economic status. CAUs with the lowest scores are among the locations in New Zealand with low access to resources and higher rates of deprivation. One of them, Wairoa, has been included in the Resilient Rural Communities programme; a discussion of that research can be found in Smith (2019). The group with the highest score includes several CAUs in the Queenstown Lakes District, which has been one of the fast-growing areas of New Zealand in recent times. Thus, the scoring created with the analysis seems to capture something about the lived experience of people and households in these rural CAUs.

Table 7 Loadings for components from PCA

	Principal component														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Access.to.Internet	0.343						0.153			0.466	-0.463	0.420	0.468		
Percent.smokers	-0.326		0.241		0.111				-0.288	-0.173	0.193	0.795	-0.171		
Unemployment	-0.308			-0.239	0.254	-0.181	0.273	-0.162	-0.262	0.664	0.199	-0.188	-0.164	0.146	
Māori.percentage	-0.305	-0.280			0.183		0.284	-0.295		-0.200			0.191	-0.708	-0.163
Tertiary	0.293	-0.325	-0.133	-0.237		-0.184				-0.136			-0.363	-0.202	0.698
Secondary	0.292	-0.357		-0.232							-0.223		-0.471		-0.656
Te.reo.percentage	-0.282	-0.358	-0.102		0.179		0.268		0.165	-0.311	-0.361		0.127	0.612	0.110
No.vehicle.percentage	-0.269			-0.576	-0.163		-0.151	0.317	-0.442	-0.139	-0.404	-0.102	0.204		
Bachelor	0.264	-0.400		-0.278						-0.145	0.584		0.510	0.135	-0.100
No.access.to.telecommunications	-0.264	-0.341					-0.264	0.580	0.521	0.292		0.117		-0.129	
Population change	0.188			-0.258	0.679	0.634	-0.160								
Religious.affiliation	-0.175		-0.433		-0.456	0.588	0.390	0.191							
Home.ownership.percentage	0.155	0.212	-0.496	0.220	0.358	-0.346	0.257	0.497	-0.188	-0.112					-0.144
Dependency.ratio	-0.142	0.311	-0.487	-0.420		-0.204	-0.185	-0.347	0.412			0.299			
Volunteering.percentage	-0.125	-0.350	-0.469	0.330			-0.601	-0.156	-0.363						

Note: Indicators have been ordered from highest to lowest loading, using the absolute value of loadings on the first principal component.

Table 8 Selected scores for CAUs based on the first principal component

CAU name	Score	CAU name	Score
<i>Lowest scores – 1st to 10th</i>			
Murupara	-0.586	Hokianga North	-0.325
Moerewa	-0.457	Levin South	-0.288
East Cape	-0.451	Kaitaia West	-0.286
Kaikohe	-0.417	Wairoa	-0.263
Cape Runaway	-0.382	Kawakawa	-0.263
<i>Middle scores – 150th to 159th</i>			
Nireaha-Tiraumea	0.280	Masterton West	0.284
Norsewood-Herbertville	0.280	Pohonui-Porewa	0.286
Oamaru South	0.281	Te Waewae	0.287
Geraldine	0.282	Papatawa	0.289
Richmond Heights	0.282	Rakaia	0.290
<i>Highest scores – 299th to 308th</i>			
Rapaura	0.564	Wanaka	0.627
Sunshine Bay	0.567	Kelvin Heights	0.705
Arrowtown	0.579	Wakatipu Basin	0.714
Queenstown Hill	0.605	Arthurs Point	0.859
Hawea	0.627	Lake Hayes South	0.951

Discussion

The PCA summarises something meaningful in the indicators

It is important at this point to review the statistical process. First, a set of indicators was identified based on the international literature on wellbeing and resilience (Fielke, Kaye-Blake, & Vibart, 2017). Next, these indicators were collected for a set of Census Area Units in New Zealand, with the CAUs identified according to two criteria (population and rural designation). A standard statistical technique, PCA, was applied to the data set, and it recovered a first principal component that captured just over half of the variation in the data set. The calculated loadings were applied to the data set to create scores for each CAU. Up to this point in the analysis, the process was essentially mathematical: we found data, isolated eigenvectors and made some calculations. The process was essentially insensitive to people's lives or the resilience of communities.

The resulting scores and ordered list, however, do seem to be meaningful. They reflect something about the lives of people and households in these places. The process seems to have created a one-dimensional socio-economic status indicator that can identify whether places are doing, for lack of nuance, 'better' or 'worse'.

This finding is important in two ways. First, it suggests that about one-half of the wellbeing or resilience of rural communities in New Zealand can be linked to standard social, economic and cultural indicators. Taken together, the indicators reflect structural conditions that these communities face. Second, the results suggest that about half the variation in the indicators is not captured in a one-dimensional ranking. As the work in Resilient Rural Communities has found, official statistics are in partial agreement with community members' perceptions of their communities' resilience (Payne, et al., 2018), but only partial agreement. This analysis tends to support that earlier finding. The results thus provide support for community-focused research. It is important to understand the context that surrounds a rural community, but it is also important to work directly with them to understand their specific experiences (Brown, Kaye-Blake, & Payne, 2019).

The next step could be to understand what this single principal component represents – and what it does not. This would be a sense-making exercise. One way to approach it would be to return to rural communities with the results of the analysis to explore them collaboratively. The discussion could focus on the score for each community or CAU: what local conditions is it reflecting, and what has been left out? Another question would be the reactions to the ratings of *other* communities: are they fair or not? The discussion could also explore the loadings or weighting for the indicators: what are rural residents' reactions to them?

The single component clearly has an economic component. Unemployment is one of the indicators and it has a fairly high loading. Education is also important for the component, and educational attainment has strong links to economic outcomes. However, the component is not just economic. It isn't just about jobs and earnings. For example, the dependency ratio (a measure of the working age population versus the young and the elderly) is one of the least important indicators. By contrast, access to the internet and the prevalence of smoking have the two highest loadings, and the percentage of people identifying as Māori is fourth (after unemployment). The single component therefore captures something more than just economic drivers.

The analysis can be a useful initial tool for researchers

The results suggest that official statistics can provide a simple indication of resilience for rural communities in New Zealand, but that this indication will not capture the full experience of these communities. This finding can be useful to researchers involved in agriculture and rural communities to provide initial guidance and help with prioritisation.

The list and the analysis provide an objective, quantitative approach to making a quick assessment of a location. Researchers who are new to an area, or a research programme that is expanding into new locations, can obtain some initial guidance and have some confidence in the assessment. They can identify similar communities from prior research and use the information to develop idea about likely resources and challenges.

The statistical assessment could also be useful for making prioritisation decisions.

- Research is constrained by resources and researchers have to make decisions about where to do their work. They may find it useful to know where a community sits on this type of relative socio-economic status scale. They may even be able to link research topics to relevant locations based on performance on specific indicators.
- The scale could also be useful for locating anomalies. These may be communities that seem to be performing significantly better or worse than their statistical profile and so would warrant further study.
- The scale and indicator loadings could suggest areas of emphasis, such as resilience dimensions that could be the focus of future research. For example, the educational attainment indicators have high loadings, which suggests that understanding the linkages between education and rural resilience could be a useful area of research.
- The analysis provides more information for people who evaluate research, either at the proposal stage or the review stage. It provides an external source of information about communities who form part of research, one that is independent from the judgment of the researchers themselves.

The statistical findings also validate the engaged approach taken by Resilient Rural Communities. They show the variation across the communities that is not captured by a simple scale. The result underscores the importance of engaging with stakeholders and communities, because the view based on official statistics may be incomplete. As the wider work has shown, resilience is complex and people have options and agency in their lives (Brown, Kaye-Blake, & Payne, 2019).

References

- Brown, M., Kaye-Blake, B., & Payne, P. (Eds.). (2019). *Heartland Strong: how rural New Zealand can change and thrive*. Palmerston North: Massey University Press.
- Fielke, S., Kaye-Blake, W., & Vibart, R. (2017). *Operationalising rural community resilience: framing indicators for measurement*. Hamilton, New Zealand: AgResearch Limited.
- Fielke, S., Kaye-Blake, W., Mackay, A., Smith, W., Rendel, J., & Dominati, E. (2018). Learning from resilience research: Findings from four projects in New Zealand. *Land Use Policy*, 70, 322-333.
- Kaye-Blake, W., Stirrat, K., Smith, M., & Fielke, S. (2017). *Testing indicators of resilience for rural communities*. Palmerston North: AgResearch.
- Kolman, B. (1986). *Elementary linear algebra* (4th ed.). New York: MacMillan Publishing Company.
- MacCannell, D. (1988). Industrial agriculture and rural community degradation. In L. Swanson (Ed.), *Agriculture and community change in the United States* (pp. 15-75). Boulder, CO: Westview Press.
- NIST/SEMATECH. (2012). *e-Handbook of Statistical Methods*. Washington, D.C.: NIST, U.S. Department of Commerce. Retrieved May 17, 2019, from <http://www.itl.nist.gov/div898/handbook/>
- Payne, P. R., Kaye-Blake, W. H., Stirrat, K. A., Ellison, R. A., Smith, M. J., & Brown, M. (2018). Identifying resilience dimensions and thresholds: evidence from four rural communities in New Zealand. *Resilience*. Retrieved from <https://doi.org/10.1080/21693293.2018.1545339>
- Rodgers, J. L., & Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42(1), 59-66.
- Smith, W. (2019). Wairoa: resilience and change. In M. Brown, B. Kaye-Blake, & P. Payne (Eds.), *Heartland Strong: how rural New Zealand can change and thrive* (pp. 32-45). Palmerston North: Massey University Press.
- Statistics New Zealand. (2004). *New Zealand: An urban/rural profile*. Wellington. Retrieved May 20, 2019, from http://archive.stats.govt.nz/browse_for_stats/Maps_and_geography/Geographic-areas/urban-rural-profile.aspx
- Statistics NZ. (nd). Area unit. Wellington. Retrieved May 01, 2019, from archive.stats.govt.nz/methods/classifications-and-standards/classification-related-stats-standards/area-unit/definition.aspx
- Statistics NZ. (nd). Statistical standard for meshblock. Wellington. Retrieved May 23, 2019, from <http://archive.stats.govt.nz/methods/classifications-and-standards/classification-related-stats-standards/meshblock/definition.aspx>
- Vyas, S., & Kumaranayake, L. (2006, November). Constructing socio-economic status indices: how to use principal components analysis. *Health Policy and Planning*, 21(6), 459-468.
- Young, R. C., & MacCannell, D. (1979). Predicting the quality of life in the United States. *Social Indicators Research*, 6, 23-40.

Appendix A Results for all CAUs

Table 9 Scores for CAUs based on first principal component

CAU name	Score	CAU name	Score
Murupara	-0.586	Playford Park	-0.032
Moerewa	-0.457	Mataura	-0.028
East Cape	-0.451	Masterton East	-0.028
Kaikohe	-0.417	Paeroa	-0.028
Cape Runaway	-0.382	Opunake	-0.012
Hokianga North	-0.325	Ahipara	-0.006
Levin South	-0.288	Trident	-0.004
Kaitaia West	-0.286	Aotea	-0.004
Wairoa	-0.263	Putaruru	0.001
Kawakawa	-0.263	Herekino	0.003
Taumarunui Central	-0.262	Otorohanga	0.005
Kaitaia East	-0.248	Edgecumbe	0.015
Opotiki	-0.233	Levin West	0.018
Kawerau	-0.213	Houhora	0.025
Hokianga South	-0.194	Eltham	0.035
Patea	-0.176	Mangapa-Matauri Bay	0.040
Matahina-Minginui	-0.173	Motutangi-Kareponia	0.042
Stanley Park	-0.166	Whakatane West	0.042
Turangi	-0.142	Pahiatua	0.050
Strathmore	-0.136	Tauhara	0.057
Maketu Community	-0.132	Waihi	0.057
Ngapuhi-Kaikou	-0.117	Waiouru	0.071
Tarrangower	-0.111	Pokere-Waihaha	0.074
Te Kuiti	-0.105	Rotoma	0.075
Shannon	-0.095	Hawera South	0.075
Urewera	-0.092	Whakatane North	0.081
Dannevirke East	-0.089	Marton	0.082
Matarawa	-0.081	Waimate	0.086
Dargaville	-0.043	Morrinsville West	0.092
Taihape	-0.037	Wellsford	0.092

CAU name	Score	CAU name	Score
Moanataiari	0.096	Solway North	0.176
Cobden	0.097	Orana Park	0.177
Tarndale-Rakauroa	0.103	Parawai	0.177
Waiotahi	0.106	Matamata North	0.177
Karikari Peninsula- Maungataniwha	0.108	Paraonui	0.178
Mokauiti	0.109	Rangiuru	0.180
Waipukurau	0.116	Ashburton Central West	0.180
Milton	0.123	Hampstead	0.181
Coromandel	0.126	Ngaumutawa	0.182
Levin East	0.127	Taipa Bay-Mangonui	0.189
Taupo Central	0.132	Picton	0.189
Dannevirke West	0.133	Kaikoura Township	0.192
Ashburton Central East	0.135	Waimataitai	0.193
Tangiwai	0.138	Runanga-Rapahoe	0.195
Ngapuke	0.140	Kahui	0.197
Parkside	0.142	Seaview	0.197
Hawera North	0.149	Marotiri	0.197
Mayfield	0.151	Hokitika Urban	0.200
Levin North	0.153	Westport Urban	0.201
Waipawa	0.156	East Gore	0.202
Oamaru North	0.160	Whangamata	0.203
Te Aroha	0.160	Paihia	0.204
Timaru Gardens	0.163	Reefton	0.213
Temuka	0.163	Te Kauwhata	0.218
Stratford West	0.165	Balclutha	0.219
Stratford East	0.166	Katikati Community	0.221
Kaipara Coastal	0.167	Matamata South	0.228
Netherby	0.171	Wharekaka	0.229
Solway South	0.173	Tiniroto	0.230
Ohakune	0.173	Motueka East	0.231

CAU name	Score	CAU name	Score
Marokopa	0.233	Norsewood-Herbertville	0.280
Redwoodtown	0.234	Oamaru South	0.281
Ngatea	0.235	Geraldine	0.282
Allandale-Mokorua	0.236	Richmond Heights	0.282
Winton	0.237	Masterton West	0.284
Kapuni	0.239	Pohonui-Porewa	0.286
Blenheim Central	0.239	Te Waewae	0.287
Lansdowne	0.239	Papatawa	0.289
Oamaru Central	0.239	Rakaia	0.290
Waihou-Walton	0.240	Hikuai	0.290
Golden Springs	0.244	Tairua	0.293
Marsden Point-Ruakaka	0.245	Amberley	0.293
Whenuakura	0.247	Whitikahu	0.296
Carterton	0.248	Whitney	0.300
Pongakawa	0.254	Morrinsville East	0.300
Turua	0.254	Watlington	0.300
North Gore	0.255	Glenwood	0.302
Maungaru	0.258	Maori Park	0.302
Otorohanga Rural West	0.259	Otorohanga Rural East	0.306
Marchwiell	0.262	Okoki-Okau	0.308
Hauraki Plains	0.268	Russell	0.309
Ohinemuri	0.268	Springdale	0.310
Waipa Valley	0.269	Alexandra	0.311
Rehia-Oneriri	0.271	Motueka West	0.312
West Gore	0.275	Lake Alice	0.313
Mangatoki-Moeroa	0.275	Fairfax	0.315
Tinwald	0.277	Toetoes	0.316
Nukuhau	0.277	Waihi Beach	0.316
Whitianga	0.279	Greymouth South	0.318
Nireaha-Tiraumea	0.280	Allenton West	0.322

CAU name	Score	CAU name	Score
Otakiri	0.324	Allenton East	0.368
Waikouaiti	0.326	Highfield	0.368
Waipu	0.326	Waihao	0.369
Mangawhai Heads	0.326	Tikokino	0.370
Maramarua	0.326	Waikaia	0.372
Toko	0.330	Oroua Downs-Waitohi	0.378
Okauia	0.331	Spring Creek-Grovetown	0.380
Mangatainoka	0.331	Ngakuru	0.381
Fraser Park	0.332	Elsthorpe-Flemington	0.382
Pembroke	0.334	Marlborough Sounds Terrestrial	0.382
Waiotira-Springfield	0.335	Twizel Community	0.383
Hilltop	0.335	Te Anau	0.385
Onewhero	0.335	Waerenga	0.385
Raurimu	0.338	Clutha	0.386
Cape Rodney	0.338	Maniototo	0.386
Tahuroa	0.339	Hurunui	0.386
Arapuni	0.341	Mangaore-Manakau	0.388
Kiwitea	0.342	Waiopehu	0.389
Pleasant Point	0.344	Bruce	0.389
Teviot	0.346	Maheno	0.391
Chertsey	0.346	Cromwell	0.393
Tuapeka	0.347	Waikawa	0.394
Kerikeri	0.348	Kaikoura Rural	0.395
Te Rerenga	0.351	Kaimata	0.395
Waituna	0.357	Mangatawhiri	0.398
Tahawai	0.357	Springlands	0.398
Awhitu	0.358	Whareama	0.399
Warkworth	0.358	Dacre	0.400
Martinborough	0.359	Snells Beach	0.400
Takaka	0.366	Hokonui	0.400

CAU name	Score	CAU name	Score
Gleniti	0.401	Mararoa River	0.464
Greytown	0.401	Mangawhai	0.464
Waianiwa	0.404	Te Uku	0.466
Chatton	0.404	Mt Somers	0.467
Riwaka	0.404	South Beach-Camersons	0.468
Hokitika Rural	0.407	Maraekakaho	0.470
Franz Josef	0.411	Ohope	0.470
Waihou Valley-Hupara	0.413	Golden Bay	0.479
Otaua	0.419	Queenstown Bay	0.481
Waipahihi	0.419	Mt Holdsworth	0.482
Amuri	0.423	Kopuaranga	0.495
Renwick	0.424	Kahutara	0.496
Hinds	0.427	Tuamarina	0.496
Orari	0.427	Methven	0.497
Malvern	0.428	Otaki Forks	0.499
Hanmer Springs	0.428	Frankton	0.501
Clyde	0.429	Kaiteriteri	0.509
Oruanui	0.436	Te Wharau	0.511
Rotorangi	0.436	Motueka Outer	0.514
Glenbrook	0.436	Ashley Gorge	0.516
Aongatete	0.438	Acacia Bay	0.520
Mackenzie	0.440	Wairau South	0.523
Levels	0.441	Cape Rodney South	0.524
Akaroa Harbour	0.445	Opaki-Fernridge	0.534
Westport Rural	0.446	Little River	0.550
Kapiro	0.453	Lake Tekapo	0.554
Witherlea	0.454	Pohangina	0.558
Selwyn-Rakaia	0.454	Dunstan	0.563
Lakewood	0.456	Rapaura	0.564
Leithfield	0.458	Sunshine Bay	0.567

CAU name	Score	CAU name	Score
Arrowtown	0.579	Kelvin Heights	0.705
Queenstown Hill	0.605	Wakatipu Basin	0.714
Hawea	0.627	Arthurs Point	0.859
Wanaka	0.627	Lake Hayes South	0.951

Appendix B Restrictions

This report has been prepared solely for the purposes stated herein and should not be relied upon for any other purpose. We accept no liability to any party should it be used for any purpose other than that for which it was prepared.

This report is strictly confidential and (save to the extent required by applicable law and/or regulation) must not be released to any third party without our express written consent which is at our sole discretion.

To the fullest extent permitted by law, PwC accepts no duty of care to any third party in connection with the provision of this report and/or any related information or explanation (together, the “Information”). Accordingly, regardless of the form of action, whether in contract, tort (including without limitation, negligence) or otherwise, and to the extent permitted by applicable law, PwC accepts no liability of any kind to any third party and disclaims all responsibility for the consequences of any third party acting or refraining to act in reliance on the Information.

We have not independently verified the accuracy of information provided to us, and have not conducted any form of audit in respect of the organisation for which work is completed. Accordingly, we express no opinion on the reliability, accuracy or completeness of the information provided to us and upon which we have relied.

The statements and opinions expressed herein have been made in good faith, and on the basis that all information relied upon is true and accurate in all material respects, and not misleading by reason of omission or otherwise.

The statements and opinions expressed in this report are based on information available as at the date of the report.

We reserve the right, but will be under no obligation, to review or amend our report, if any additional information, which was in existence on the date of this report, was not brought to our attention, or subsequently comes to light.

This report is issued pursuant to the terms and conditions set out in our engagement letter and the Terms of Business dated 26 October 2018.